REVIEW REPORT

Journal: HESS
Paper: hess-2024-335
Title: Return period of high-dimensional compound events. Part II: Analysis of spatially-variable precipitation
Author(s): Manuel Del Jesus, Diego Urrea Méndez, and Dina Vanessa Gomez Rave

GENERAL COMMENT.
For the reasons given below, I recommend a **REJECTION (WITH POSSIBLE RESUBMISSION).**

SPECIFIC COMMENTS.

Line(s) 26–27

AUTHOR(s). Depending on the hydrological scale used, compound precipitation events measured at different rain gauges may contain a considerable number of zeros (zero-inflated data).

REFEREE. From a Statistical point of view, it is a nightmare, and the problem is far from being resolved. The introduction of a mixed model, and then considering only positive rainfall values, does not seem to fix the question, for introduces other problems (see my comment below).

Line(s) 36–38

AUTHOR(s). One of the first mixed models applied in a bivariate approach was developed by (Shimizu, 1993). This approach represents a copula-based mixed distribution function composed of a continuous part (observations greater than zero) and a discrete part (observations at zero).

REFEREE. The model by Shimizu (1993) is not copula-based: the modeling via copulas is present nowhere in the paper, also considering that it deals with a mixture of discrete-continuous distributions. Incidentally, none of the marginals used in that paper is heavy-tailed, possibly inadequate to deal with rainfall extremes.

Line(s) 43–44

AUTHOR(s). Another fundamental definition when discussing floods corresponds to the notion of the return period (RP). Salvadori et al. (2011) defines the RP as the time elapsed between two successive occurrences of a prescribed event.

REFEREE. The true novelty of the paper by Salvadori et al. (2011), beyond the mathematical formalization of a multivariate notion of RP, and the introduction of original multivariate design techniques, is that the calculation of the RP is written in terms of copulas only in any multi-dimensional setting (not only bivariate).

Line(s) 56–57

AUTHOR(s). In this context, this study has two main objectives: (I) to expand the methodological framework for modeling data with zero intermittency from a bivariate (Shimizu, 1993; Serinaldi, 2008; Villarini et al., 2008), to a five-dimensional approach…

REFEREE. Is 5 a magic number, in hydrology or elsewhere? Increasing the dimension is not a synonymous of novelty: what about if tomorrow I publish a paper on a 6-dimensional approach? It may solve problems in 6 dimensions, but may not change the paradigms… in addition, you only dealt with a sub-class (Group 32) of the 5 dimensional problem…

Line(s) 72–73

AUTHOR(s). This consideration leads to the incorporation of multivariate mixed models, which will be detailed in this chapter.

REFEREE. The mixed model ignores/spoils the correlation structure of the sequences of $(0, > 0)$'s in the rainfall time series, and in turn the COMPOUND nature/feature of the events.

For example, the total precipitation in the two sequences A and B below is the same (0 means no rain, 1 means rain), but the COMPOUND impact of series B could be devastating as compared to the one of series A:

A=[0,1,0,1,0,1,0,1,0,1]

B=[1,1,1,1,1,0,0,0,0,0]

Here, the correct approach would be to use a stochastic renewal process, as in

G. Salvadori and C. De Michele. Statistical characterization of temporal structure of storms. Advances in Water Resources, 29(6):827–842, 2006. doi: 10.1016/j.advwatres.2005.07.013

Line(s) 77

AUTHOR(s). 4. Hazard scenario

REFEREE. A reference to

Salvadori, G., Durante, F., De Michele, C., Bernardi, M., and Petrella, L.: A Multivariate Copula-Based Framework for Dealing with Hazard Scenarios and Failure Probabilities, Water Resources Research, 52, 3701–3721, https://onlinelibrary.wiley.com/doi/abs/10.1002/2015WR017225, 2016

should be put here: it is the first paper where Hazard Scenarios are first formalized in terms of Copulas, including those mentioned by the Authors.

Line(s) 80

AUTHOR(s). Specifically, we compute the critical surface…

REFEREE. Surface or hyper-surface, with dimension larger than 2?

Line(s) 94–96

AUTHOR(s). To ensure regional representativeness, "regional events" are then selected based on simultaneous rainfall contribution across all stations and, for non-independent events, the highest total precipitation.

REFEREE. This sentence is quite obscure: what do you mean by "for non-independent events"? What are the events you are considering? Monthly maxima at different stations? What kind of independence do you consider? Spatial-Pairwise? Spatial-Global? Note that they are different: Pairwise independence may not imply Global one… How do you test it? How do you identify a "homogenous" region? Via clustering procedures?

Line(s) 105 & Eq. (1)

AUTHOR(s). In this context, we have extended the model proposed by Serinaldi (2008)…

REFEREE. I do not think it is an extension, except perhaps for the dimension, but then it would be trivial: you simply try to account for the probability of mutually exclusive events in dimensions larger than 2, nothing too special that could justify a specific paper…

Φ is not precisely defined, what should it represent? A joint CDF? Eq. (1) looks like a linear combination of probabilities: what is its meaning? What is the domain/support of the parameters p's, and their inter-relationships? No information/explanation is provided…

Furthermore, written in this way, the formula may account twice for the probability of zero rainfall: in fact, by definition, $F(x) = P(X \leq x)$ (in whatever form you write it, for one or more variables) includes the case that the variable(s) take on the value 0, even if the threshold x is strictly larger than 0: the probabilities in Eq. (1) look more like conditional ones. In addition, P_0 should depend upon the location, I would be surprised if it were the same at all stations.

In all cases, you should prove that Φ is a genuine probability distribution, which however suffers from over-parametrization (i.e., the number of p's): estimating all these parameters in a high-dimensional space is a torment, at least from a numerical standpoint, for the estimates almost certainly never correspond to optimal values (at best, they are suboptimal in the most favorable cases).

Line(s) 125–126

AUTHOR(s). Gaussian Copula without intermittency (Gaussian): This approach considers the joint dependence between compound rainfall events without accounting for zero intermittency, including all rainfall data without exception.

REFEREE. The presence of 0's yields Ties, which adversely affect (spoil) statistical techniques: how do you manage such a problem, given the fact that no effective solutions are present in Literature? The failure of the Gaussian approach may be due to the fact that, as remarked below, it is inadequate for dealing with extremes, but also to the fact that Ties play against it in this approach: you must make things clear.

Line(s) 128–129

AUTHOR(s). This method models each group using the Gaussian copula, leveraging its ability to model high dimensions.

REFEREE. Unfortunately, a Gaussian framework is unsuitable for dealing with maxima, such as those considered in this paper…

Line(s) 130–132

AUTHOR(s). This approach utilizes R-vine structures with Gaussian copulas to model all pairs of series. It combines the flexibility of R-vines for capturing complex dependence structures with the efficiency of Gaussian copulas for pairwise modeling.

REFEREE. I am not sure that a Gaussian copula could be "efficient" (whatever you mean with this unspecified feature) if the true dependence structure is not Gaussian itself. A Gaussian copula has feasible mathematical properties, but also strong limitations, especially considering Extreme phenomena, as abundantly pointed out in Literature.

Line(s) 133

AUTHOR(s). Vine extreme copulas (Vine extreme): This approach uses R-vine structures with a diverse set of bivariate copulas…

REFEREE. Perhaps, dealing with maxima, Extreme Value copulas should better be used, but these exclude the case of negative dependencies: a justification is required here.

Line(s) 143–146

AUTHOR(s). The selection of the Gaussian copula […] is supported by its frequent application in climate and hydrological research focused on simulating extreme conditions (Chen and Guo, 2019).

REFEREE. This sentence/explanation sounds like a suicide, for it reads as: since (inexperienced) practitioners frequently use the Gaussian copula, this justifies its use, and therefore we use it. No comment.

Line(s) 149–150

AUTHOR(s). To carry out our analysis comprehensively, we have selected 5 strategically distributed rain gauge stations…

REFEREE. What do you mean by “strategic”? Do you mean “representative” of the hydrological regime (whatever the word “representative” could mean)? What regionalization/clustering procedures/criteria did you use to decide that these are “strategic”? Or these stations are the only ones available (and so the number 5 has a justification)?

Line(s) 155–156

AUTHOR(s). A rigorous quality control process was implemented, including outlier identification (Gonzalez and Bech, 2017), review of repeated values…

REFEREE. What do you mean by “review of repeated values”? And rigorous with respect to what benchmarks?

Line(s) 175–177

AUTHOR(s). Given the considerable number of groups and to simplify the interpretation of the findings, we will focus on the group where rainfall occurs simultaneously in all stations (group 32 - Fig. 1).

REFEREE. So what? You introduce a tricky model, then you realize it is too complex, and thus you use the simplest case given by Group 32: essentially, it corresponds to a classical “AND” hazard scenario. The fact that the model is a mathematical mess was already clear in Eq. (1), so why not considering the case of Group 32 directly, which simplifies the discussion, as well as the mathematical treatment. In practice, you boasted about solving a problem in 5 dimensions, but then you only dealt with a specific sub-case.

Line(s) 185–187

AUTHOR(s). Figure 5 presents the autocorrelation plots calculated for group 32. When analyzing the autocorrelation plot of the five event series, it is observed that there is no significant correlation between values at different time intervals.

REFEREE. To the best of my understanding of the plot, quite a few estimates of the ACF are outside a (traditional) 5% Confidence Band, and thus I would suspect that the data ARE auto-correlated…

Line(s) 195

AUTHOR(s). In the upper triangular matrix, Kendall’s τ values are displayed in a heatmap…

REFEREE. Confidence Intervals for the estimates must also be provided.

Line(s) 203–204

AUTHOR(s). The results from this indicator showed that both upper and lower tail dependence were present in the data.

REFEREE. Believe me, with such data you cannot really claim anything about the possible (statistical) presence of Tail Dependence: this is just visual statistics, too often a deceiving practice used by inexperienced practitioners…

Line(s) 226–227

AUTHOR(s). Additionally, the QQ plots for each group were checked, and it was observed that the GEV adequately represented the tail behavior.

REFEREE. Formal Monte Carlo Goodness-of-Fit tests, and the corresponding p-values, would be less visual and more objective (e.g., Kolmogorov-Smirnov, or even better Anderson-Darling ones).

Table 1

REFEREE. In Table 1, GoF p-values are missing for the first two cases, they should be shown.

Line(s) 256–257

AUTHOR(s). To analyze the results for the remaining groups, a box plot was constructed, as presented in Fig. 8, where the distributions of AIC for all groups in each proposed approach are compared.

REFEREE. You must first check that the model is admissible via a GoF test, and then (and only then) select the "best" model (according to some criterion) ONLY among the admissible ones. The plots of the AIC's alone in Fig. 8 are of little interest/significance: the corresponding models could all be non-admissible without, first, carrying out suitable GoF tests.

Line(s) 271–272

AUTHOR(s). The first was crucial for assessing whether the dependency of the observed values was maintained, reduced, or improved.

REFEREE. How you could "improve" a dependence remains a mystery to me...

Line(s) 283–284

AUTHOR(s). This finding supports the ability of the copulas used to accurately capture and reproduce the behavior of the real variables in terms of their extremes and dependencies.

REFEREE. Statistically speaking, at most you can hope it: your conclusions are only based on visual analyses, be careful.

Line(s) 291

AUTHOR(s). Based on the analyzed results, the Vine extreme approach demonstrated its ability to reproduce upper tail dependencies.

REFEREE. You should add: assuming it is really present.

Line(s) 324–327

AUTHOR(s). To calculate the critical level t, it was necessary to calculate the 100-year RP. Considering that we have more values per year than in the case of annual maximum, the quantiles in this case move to the extreme part of the distribution. Note also that each Kendall function is calculated from the continuous part of the function described in Eq. (1), that is, it considers the complete CDF.

REFEREE. This claim is quite obscure, and should be clarified. Intuitively, it should be enough to properly set the constant μ in the definition of the Kendall RP to fix the right temporal scale (e.g., $\mu = 1/12$). However, this looks like a Kendall RP conditional to the fact that rain is present.

Line(s) 329–330

AUTHOR(s). The best-performing approach (4) obtained a critical value of 0.993, while the lowest-performing approach (1) obtained a critical value of 0.778.

REFEREE. Assuming that these results make sense, you should interpret them, and discuss the consequences.

Line(s) 341–344

AUTHOR(s). This procedure was iterated until we obtained a sufficient number of events on the critical layer for each approach. Iteration was necessary because, given the specific nature of the critical level t, only a small fraction of the synthetic events would correspond exactly to this value.

REFEREE. Frankly speaking, I really doubt that any of the events generated actually lies on the critical layer, if only for the sake of numerical approximation. Most likely, you have fixed some tolerance coefficient: you must clearly explain how you accept that an event lies on the critical layer.

Line(s) 354–355

AUTHOR(s). Solving this loss of precision in high dimensions was easy because we had sufficient event combinations on the critical layer. For each combination of events, we calculated the density function and selected the one with the highest density.

REFEREE. It is not clear what you mean by a “combination of events”, and how it is chosen. What is its sample size and how is it decided? What is its density function (the joint one?) More details must be given for the sake of discussion and reproducibility.

Line(s) 371–372

AUTHOR(s). To compare the results of univariate and multivariate analysis, it was necessary to calculate the average precipitation in the watershed using both approaches.

REFEREE. Average precipitation could have little to do with the Extreme Value approach: I understand that it is part of common hydrological practice, but then it seems that the Authors are playing at the same time on two different layers, as if they were trying to have a foot in both camps. A justification is required here.

Line(s) 379–381

AUTHOR(s). Compared to this method, the Gaussian, Gaussian Groups, and Vine Gaussian models tend to underestimate the events, while the Vine t-student overestimates them.

REFEREE. Here, as well as below, you cannot speak about under- or over-estimates: this makes sense only if you know the true value. Here you can only speak about relative smaller/larger values.

Line(s) 399–ff.

REFEREE. The Discussion and the Conclusions sections could/should be merged in a single section “Discussion & Conclusions”.